**Listing of Claims:**

1-75. (Cancelled)

76. (Currently amended) A method for identifying nucleotides for variation in nucleic acids encoding a protein variant library in order to affect a desired activity, said method comprising:

(a) receiving data characterizing a training set of a protein variant library, wherein the data comprises activity and a nucleotide sequence for each protein variant in the training set;

(b) from the data, developing a computational algorithmic sequence activity model that predicts activity as a function of nucleotide types and corresponding position in the nucleotide sequence;

(c) using the sequence activity model to rank positions in a reference nucleotide sequence and/or nucleotide types at specific positions in the reference nucleotide sequence in order of impact on the desired activity;

(d) using the ranking to identify one or more nucleotides, in the reference nucleotide sequence, that are to be varied or fixed in order to impact the desired activity; ~~and~~

(e) generating a new protein variant library containing one or more of the protein variants ~~encoded by the reference nucleotide sequence~~ in which the identified nucleotides are varied or fixed in order to impact the desired activity;

(f) assaying the new protein variant library to provide activity information used to develop a new computational algorithmic sequence activity model; and

(g) using the new computational algorithmic sequence activity model to identify one or more nucleotides in a new reference nucleotide sequence that are to be varied or fixed in order to impact the desired activity.

77. (Previously presented) The method of claim 76, wherein the nucleotides to be varied are codons encoding particular amino acids.

78. (Previously presented) The method of claim 77, wherein the activity is a function of expression of nucleic acids.

79. (Currently Amended) A computer program product comprising a machine readable medium on which is provided program instructions for identifying nucleotides for variation in nucleic acids encoding a protein variant library in order to affect a desired activity, said instructions comprising:

(a) code for receiving data characterizing a training set of a protein variant library, wherein the data comprises activity and a nucleotide sequence for each protein variant in the training set;

(b) code for developing a computational algorithmic sequence activity model from the data, which sequence activity model predicts activity as a function of nucleotide types and corresponding position in the nucleotide sequence;

(c) code for using the sequence activity model to rank positions in a reference nucleotide sequence and/or nucleotide types at specific positions in the reference nucleotide sequence in order of impact on the desired activity;

(d) code for generating a ranked list of the nucleotide positions and/or the nucleotide types at specific positions in the reference nucleotide sequence; ~~and~~

(e) code for using the ranking to identify one or more nucleotides, in the reference nucleotide sequence, that are to be varied or fixed in order to impact the desired activity;

(f) code for receiving activity data characterizing a new protein variant library containing one or more protein variants having sequences in which the identified nucleotides were varied or fixed in order to impact the desired activity;

(g) code for using the activity data characterizing a new protein variant library to develop a new computational algorithmic sequence activity model;

(h) code for using the new computational algorithmic sequence activity model to identify one or more nucleotides in a new reference nucleotide sequence that are to be varied or fixed in order to impact the desired activity; and

(i) code for outputting information, in a user readable format, identifying members of the new protein variant library.

80. (Previously presented) The computer program product of claim 79, wherein the nucleotides to be varied are codons encoding particular amino acids.

81. (Previously presented) The computer program product of claim 79, wherein the activity is a function of expression of nucleic acids.

82.-100. (Cancelled)

101. (Withdrawn) The method of claim 76, wherein (e) comprises generating a new protein variant library wherein the sequences of the members of the new protein variant library comprise amino acid residues encoded by the identified nucleotides varied or fixed in order to impact the desired activity.

102. (Withdrawn) The method of claim 101, wherein (e) comprises expressing the new protein variant library from polynucleotides encoding members of the new protein variant library and wherein the polynucleotides are prepared by gene synthesis.

103. (Withdrawn) The method of claim 101, wherein (e) comprises expressing the new protein variant library from polynucleotides encoding members of the new protein variant library and wherein the polynucleotides are prepared by mutagenesis.

104. (Withdrawn) The method of claim 101, wherein (e) comprises expressing the new protein variant library from polynucleotides encoding members of the new protein variant library and wherein the polynucleotides are prepared by performing a recombination-based diversity generation mechanism.

105. (Withdrawn) The method of claim 101, further comprising screening the new protein variant library to identify protein variants having the desired activity.

106. (Withdrawn) The method of claim 105, further comprising sequencing the identified protein variants having the desired activity.

107. (Withdrawn) The method of claim 106, further comprising repeating (a) – (c) using the activity and sequence data from protein variants in the new protein variant library.

108. (Withdrawn) The method of claim 101, wherein the members of the new protein variant library comprise the same amino acid sequence encoded by different nucleotide sequences.

109. (Withdrawn) A method for identifying nucleotides for variation in a nucleotide sequence in order to optimize the expression properties of the nucleotide sequence, said method comprising:

(a) receiving data characterizing a training set comprising a nucleotide sequence and a corresponding quantity of protein expressed for each different nucleotide sequence in the training set;

(b) from the data, developing a computational algorithmic model that predicts a quantity of protein expressed as a function of nucleotide types and corresponding position in the nucleotide sequence;

(c) using the model to rank positions in a reference nucleotide sequence and/or nucleotide types at specific positions in the reference nucleotide sequence in order of impact on the quantity of protein expressed;

(d) using the ranking to identify one or more nucleotides, in the reference nucleotide sequence, that are to be varied or fixed in order to impact the quantity of protein expressed; and

(e) expressing protein from a modified version of the reference nucleotide sequence in which the identified nucleotides are varied or fixed in order to impact the quantity of protein expressed.

110. (Withdrawn) The method of claim 109, wherein some of the nucleotides to be varied comprise codons encoding particular amino acids.

111. (Withdrawn) The method of claim 109, wherein (e) comprises expressing protein from a plurality of polynucleotides corresponding to modified versions of the reference nucleotide sequence in which the identified nucleotides are varied or fixed in order to impact the quantity of protein expressed.

112. (Withdrawn) The method of claim 111, wherein in (e) the polynucleotides are prepared by gene synthesis.

113. (Withdrawn) The method of claim 111, wherein in (e) the polynucleotides are prepared by mutagenesis.

114. (Withdrawn) The method of claim 111, wherein in (e) the polynucleotides are prepared by performing a recombination-based diversity generation mechanism.

115. (Withdrawn) The method of claim 111, further comprising determining whether the modified versions of the reference nucleotide sequence impact the quantity of protein expressed.

116. (Withdrawn) The method of claim 115, further comprising repeating (a) – (c) using the quantity of protein expressed and sequence data corresponding to the modified version of the reference nucleotide sequence.

117. (Withdrawn) The method of claim 111, wherein the modified versions of the reference nucleotide sequence encode the same amino acid sequence.

118. (Withdrawn) A computer program product comprising a machine readable medium on which is provided program instructions for identifying nucleotides for variation in a nucleotide sequence in order to optimize the expression properties of the nucleotide sequence, said instructions comprising:

(a) code for receiving data characterizing a training set comprising a nucleotide sequence and a corresponding quantity of protein expressed for each different nucleotide sequence in the training set;

(b) code for developing a computational algorithmic model from the data, which model predicts a quantity of protein expressed as a function of nucleotide types and corresponding position in the nucleotide sequence;

(c) code for using the sequence activity model to rank positions in a reference nucleotide sequence and/or nucleotide types at specific positions in the reference nucleotide sequence in order of impact on the quantity of protein expressed;

(d) code for generating a ranked list of the nucleotide positions and/or the nucleotide types at specific positions in the reference nucleotide sequence; and

(e) code for using the ranking to identify one or more nucleotides, in the reference nucleotide sequence, that are to be varied or fixed in order to impact the quantity of protein expressed.

119. (Withdrawn) The computer program product of claim 118, wherein some of the nucleotides to be varied are codons encoding particular amino acids.